AFOSR 69 - 0912 TR

AD 688227

A Graph Theoretical Approach to Clustering

Barbara Ann Crow

Department of Psychology
University of Kansas
Lawrence, Kansas

Technical Report No. 1

October 1968

MODERNIZATION AND RAPID ACCULTURATION PROJECT

University of Kansas

1. This document has been approved for public release and sale ; its distribution is unlimited.

71

The author gratefully acknowledges the patient guidance of Professor Maynard W. Shelly in the preparation of this report which was submitted to the Department of Psychology in partial fulfillment of the Master of Arts Degree. The assistance of Jeff Bangert and the University of Kansas Computation Center is also appreciated.

This research was supported by the Advanced Research Projects Agency (ARPA Order No. 1071) of the Department of Defense and was monitored by the Air Force Office of Scientific Research under Contract No. F44620-68-C-0081.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

Statement of the Problem

The task of data analysis includes much more than discovering statistical significance among relationships. There is a need for methods which analyze a set of variables and interrelationships as a system -- to study the structure, configurations, and patterns within the complex interrelated whole. The computer program CLUSTR described in this paper will be part of a system of programs designed to study a set of variables as a system. Explaining the purpose of the programs, called the DIGRAF system, will help to clarify the purpose of CLUSTR.

The social scientist who studies behavior in the field must study it as part of a usually somewhat stable system which includes behaviors and parts of the environment. Behavior is maintained by a complex set of behaviors interacting with parts of the environment. It makes sense, then, to study behavior as part of a system and to analyze the interrelationships within the system or part of the system.

When a system does not remain somewhat stable over time it ceases to exist and is replaced with another system of behaviors and environmental characteristics. Social change, which is by definition a change in behavior, reflects some sort of instability in the system or parts of the system since behavior changes when behaviors and elements of the environment assume new relationships. The way to locate the sources of instability in the system, and thus to understand change if this analysis is correct, is to analyze the old and new relationships as systems (Shelly, 1968).

The author will not be concerned in this paper with defining a system or discovering what is important in a system. The DIGRAF system is to be used to analyze the set of data after a system has been defined and appropriate measures taken.

The program CLUSTR is a clustering technique based, as is the rest of DIGRAF, on graph theory. Clustering is one of the first steps in the description of a system because it defines domains of mutual influence, or groups of variables which vary together. Graph theory provides a way of taking into account direction of influence, which is not considered in other methods of clustering. A brief review of clustering techniques and graph theory will serve to introduce CLUSTR.

Clustering techniques

Although some methods of clustering existed before computers, they were limited to small numbers of variables because of the large amount of computation required. Since 1960 and the advent of relatively inexpensive computers, a whole new body of computer-related clustering techniques have been developed which attempt to group sets of data patterns into subsets which are as much "alike" as possible. Data can be clustered according to variables or according to observations (classes, individuals, etc.) and most of the techniques can be used with varying degrees of success to cluster either. A cluster is a set of data patterns (variables or observations) which are closer together according to some criteria than they are to some other set of patterns. In the case of variables (the clustering of which is one of the first steps in the analysis of a system), a cluster is a set of variables in which a change in one member affects or has an appreciable probability of affecting some of the other members of the set and has a very small probability of affecting the members of another set of variables.

A cluster of variables, in other words, defines a domain of influence within the system or subsystem. Ideally a cluster will be unique if the data is indeed clustered, but no stable clusters will be found if the data is uniformly placed in the data space. Clustering is usually distinguished from factor and principle component analysis in that clustering deals with local regions of the multidimensional data space rather than with projections of the data onto a line or plane. Clustering techniques generally are quite complex and require a great deal of computation which is possible only with the use of computers, but once programmed they can be effectively utilized by persons having relatively little mathematical and statistical training. Although some clustering techniques are not adequate to determine significance of the data in the statistical sense, they nevertheless provide descriptions of the data which are highly suggestive of new experiments and new interpretations which can lead to theory building. There is a great variety of clustering techniques, many of which are still in the process of development, and very little work has been done on comparing and evaluating the techniques and on developing methods for interpreting and examining the resulting clusters; however, the area is developing rapidly. A few types of techniques will be briefly described.

One group of clustering techniques, for instance those of Ball and Hall (1964), Tyron and Bailey (1966), and Bangert (1968), is characterized by sorting variables according to the minimum distance or maximum correlation between the variables and an arbitrarily chosen set of "cluster points". The assignment of variables to clusters as well as the position of the cluster points are improved by iteration until the centers of the clusters adequately describe the data according to some criteria.

In another group of programs (Sokal and Sneath, 1963, for example), the closest single pair of patterns is selected as a nucleus for a cluster and other variables (or observations) are assigned to that cluster on the basis of their closeness to the pair or the mean of the pair. The process is repeated for other clusters. This type of clustering is especially useful for taxonomic problems (clustering individuals).

Other techniques are based on an arbitrary partitioning of the data which is then improved by switching variables (or observations) until the best partition is formed. Another way to cluster is to decompose the distribution of the data into separate normal distributions.

None of the techniques of clustering known to the author, including a few based on graph theory, are directional; that is, none of them take into account the direction of influence between variables -- a consideration which is of vital importance in the analysis of a system. Directional influences do exist in a system and they must be taken into account. CLUSTR, a technique based on graph theory, is one possible method which can take direction of influence into account.

Graph theory

A branch of mathematics known as the theory of graphs was brought to the attention of social scientists in 1953 by two mathematicians who felt that a knowledge of the mathematics of abstract structures would be of value to investigators interested in various kinds of empirical structures (Harary and Norman, 1953). It was created by Euler in 1736 but remained an isolated contribution until the middle 1800's when interest was revived by the four color map conjecture (that only four colors were necessary to color a map so that no two adjacent nations were the same color). Until recently the major reference for graph theory was a German book (Konig, 1936). Development has been much more extensive in areas outside the U. S., especially in Eastern Europe. The first major reference in English appeared in 1962 (Ore) and the first specifically oriented to social science in 1965 (Harary, Norman, and Cartwright). Since its introduction into the social sciences, graph theory has proven to be a useful model for investigators attempting to handle such patterns of relationships as those involved in communication networks, group structure, balance theory, and sociometric choice. Graph theory can also be applied to the description of a system.

Graph theory is concerned with the abstract notion of structure as a system of points and directed lines. In the description of an empirical system of behaviors and environment we are also interested in structure; that is, we are interested in the form of the patterns of relationships which exist in the system. The direction of lines is important in graph theory, as is the direction of influence in an empirical system. A well defined concept in graph theory is that of a type of structure called a component. The components, when interpreted within a system, represent a natural grouping of variables based on the patterns of influence in the system. These components, which fulfill the basic requirement of clustering that members of a cluster covary with each other and not with members of another cluster and which take into account direction of influence, form the clusters of the program CLUSTR.

CHAPTER II: SOME ELEMENTS OF GRAPH THEORY

Definitions

In order to describe CLUSTR it is necessary to introduce some basic definitions from the mathematical theory of graphs. A more complete treatment of this theory may be found in Structural Models: An Introduction to the Theory of Graphs (Harary, Norman, Cartwright, 1965), from which the terminology of this paper is taken.

A graph is a finite collection of points, or vertices, together with a prescribed set of lines joining certain pairs of distinct points. If the lines of the graph have direction -- that is, if a and b are distinct points of a graph and the line ab from a to b is not the same as the line ba from b to a -- then the graph is called a directed graph or digraph for short. A directed path joining the points a and e is a collection of lines of the form ab, bc,...de which all go in the same direction. If the collection of lines join the same points but are not all in the same direction, such as ab, cb, ..., de, the collection is called a semipath from a to e. There may be more than one path or semipath from one point to another. A point e is said to be reachable from another point a if there is at least one path from a to e. In Figure 1 the point w is reachable from the point u along the path xv, vw, wu. The point x, however, is not reachable from the point u because they are connected only by a semipath uv, wv.

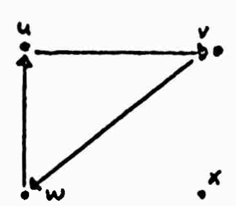


Figure 1

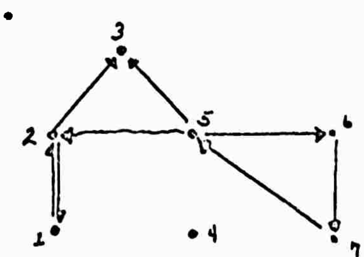


Figure 2

A digraph (or subgraph) is said to be strongly connected if every two points are mutually reachable; that is, if any point can be reached from any other point by following the arrows. A maximal strongly connected subgraph is called a strong component of the digraph. The points u, v, and w of Figure 1 together with the lines uv, uw, and wu make up a strong component (the only strong component of that digraph). Likewise, a digraph is said to be weakly connected if every two points are joined at least by a semipath; that is, every point can be reached if the direction of the lines is ignored. A maximal weakly connected subgraph is called a weak component of the digraph. The digraph of Figure 1 constitutes a weak component. Notice that a strong component is a subset of a weak component. A point or line can belong to only one strong component and to only one weak component. The components of a digraph, are then disjoint sets and no two components of the same type will overlap. For another example, the digraph of Figure 2 consists of two strong components; one consisting of the points 1 and 2 and the other with the points 5, 6, and 7. The points 1, 2, 3, 5, 6, and 7 belong to the weak component. The discrete point 4 is trivially considered a strong component (and thus a weak component) because each point is said to be reachable from itself.

Empirical meaning of a cluster

If we think of the points of a digraph as corresponding to a set of variables and the lines between points as corresponding to a measured influence greater than a given magnitude of one variable on another (the sign of the influence is not considered), then the properties of a digraph reflect the structural properties of the set of interrelationships among the variables. A line from point a to point b, in a digraph representing a set of variables and relationships, implies that variable a influences variable

b. If point e is reachable from point a then variable a directly or indirectly influences variable e. If several points are strongly connected, the corresp nding variables mutually influence each other.

The strong and weak components of a graph representing a set of variables and their interrelationships indicate the existence of clusters which have properties that can be derived by analogy from the components of a digraph. The variables belonging to a strong component vary together either in the same direction or in opposite directions. Each effects every other (directly or indirectly) and each can be expected to have some predictive value for every other. The variables belonging to the same weak component are all members of interconnected chains of influence, but any one variable cannot necessarily effect every other variable in the weak component. Any two weak components at any given cutoff level of influence are entirely independent of each other. The variables in a strong component may influence the variables in another strong component, but the influence cannot be reciprocal. If strong components are linked in this way they belong to the same weak component. Figure 3 shows strong components A and B along with the point c which are members of the same weak component which is independent of the weak component containing the strong components D, E, and F.

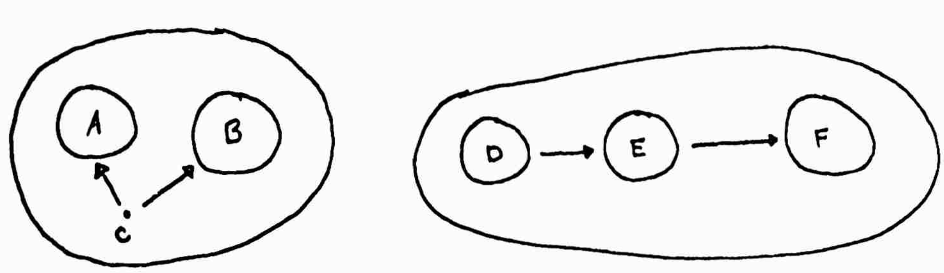


Figure 3

If a measured amount of influence above a certain level (or below in the case of the negative correlation) exists between two variables but the direction of the influence is uncertain, then it is assumed the influence occurs in both directions. If the variables represented by a digraph all have about the same influence on each other, then either the variables will all belong to the same strong component or else there will be no components, depending on the cutoff level of influence chosen. Either of these outcomes is intuitively satisfying since no clusters really exist in the data.

Matrix representation of a digraph

Computation of the strong and weak components by computers is made possible by the representation of the digraph in matrix form. The matrix corresponding to a digraph (called an adjacency matrix) is defined as a matrix whose (i,j)th entry is 1 if there is a line from point i to point j, and 0 otherwise. The diagonal elements are zero. The adjacency matrix corresponding to Figure 2 is

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Notice that the number of entries of 1 in the adjacency matrix equals the number of lines in the digraph and A is symmetric if and only if for every line from k to f there is a line from j to i.

A number of matrices can be derived from an adjacency matrix; one of these is a reachability matrix. The (i,j)th entry of a reachability matrix is 1 if j is reachable from i, and 0 otherwise. The diagonal elements are 1 since a point is said to be reachable from itself. The reachability matrix of Figure 2 is

$$R = \begin{array}{c|c} & \begin{matrix}1&2&3&4&5&6&7\end{matrix} \\ \begin{matrix}1\\2\\3\\4\\5\\6\\7\end{matrix} & \begin{bmatrix}1&1&1&0&0&0&0\\1&1&1&0&0&0&0\\0&0&1&0&0&0&0\\0&0&0&1&0&0&0\\1&1&1&0&1&1&1\\1&1&1&0&1&1&1\\1&1&1&0&1&1&1\end{bmatrix}\end{array}$$

The reachability matrix R can be directly computed from A by taking the Boolean sum of successive powers of A until taking another sum gives the same matrix. The computation of the example reachability matrix as well as an explanation of tho matrix and Boolean algebra involved may be found in Appendix A.

Transposing a matrix (interchanging the rows and columns so that the (i,j)th entry in the original matrix becomes the (j,i)th entry in the transposed matrix) has the same effect as changing the direction of the lines in the original digraph, or changing the direction of the paths in a reachability matrix. The transpose R' of our reachability matrix R is

$$R' = \begin{array}{c|c} & \begin{matrix}1&2&3&4&5&6&7\end{matrix} \\ \begin{matrix}1\\2\\3\\4\\5\\6\\7\end{matrix} & \begin{bmatrix}1&1&0&0&1&1&1\\1&1&0&0&1&1&1\\1&1&1&0&1&1&1\\0&0&0&1&0&0&0\\0&0&0&0&1&1&1\\0&0&0&0&1&1&1\\0&0&0&0&1&1&1\end{bmatrix}\end{array}$$

Taking the cross product of two matrices is defined as multiplying the corresponding (i,j)th elements of each matrix to form the (i,j)th element of the product. Since we are dealing only with entries of 0 and 1, an entry of 1 in the product indicates that the corresponding elements in the multiplicand were also 1. The cross products of R and R' has entries of 1 if and only if the corresponding points are mutually reachable. The entries of 1 in a given row (or column) are all members of the same strong component. The strong components are circled.

$$
RxR' = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}
$$

Points 1 and 2 form one nontrivial strong component and variables 5, 6, and 7 form another.

To compute the weak components of the example digraph, recall that the direction of the lines is not important in a weak component. The first thing to do, then, is to symmetricize the digraph; that is, whenever there is a line from i to j, put in also a line from j to i. Symmetricizing is easily accomplished by adding A and A'.

$$
A+A' = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}
+ \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}
= \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}
$$

The reachability matrix of the symmetricized digraph gives those points which are reachable if the direction of the lines is disregarded -- those points which make up the weak components of the digraph. The entries of 1 in a given row (or column) give the points which are members of the same strong component. The example digraph has one weak component consisting of every point except point 4. Rearranging the order of the points makes the component easier to see.

$$
R(A+A') = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}
= \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 5 \\ 6 \\ 7 \\ 4 \end{array}
\begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}
$$

CHAPTER III: THE PROGRAM CLUSTR

The program CLUSTR computes the strong and weak components of a digraph exactly as they were computed in Chapter II. The relevant programs, modified to be used separately, are included in Appendix B.

Input to CLUSTR is an adjacency matrix based on a similarity matrix. The level of cutoff above which a measure of similarity is considered significant is the only parameter of the clustering technique. Usually several adjacency matrices are computed with different values of the parameter and the resulting clusters compared. In actual practice, the similarity matrix on which an adjacency matrix is based is nondirectional and the relationships will be represented in the adjacency matrix as if the variables mutually influence each other unless some additional information can eliminate the possibility of influence in one direction. Some variables such as sex and age are nearly always independent, for others there is sound evidence from past studies to indicate the direction of the relationship, and measurements of some occur before the measurements of other variables in time. A program for computing adjacency matrices which also takes direction of influence into account is included in the DIGRAF system.

A mainline calling program reads in the parameter cards and adjacency matrix before calling CLUSTR. The first parameter card contains the number of variables in columns 1 to 5, right justified; the second card contains any titling information desired; and the third is a variable format card for the input adjacency matrix. These cards go directly before the adjacency matrix deck which is followed by either another set of parameter cards and deck or, if no more decks are to follow, a blank card. CLUSTR then computes the strong components of the digraph, prints them out, and then computes and prints the weak components. The subroutine REPUN is used to calculate the reachability matrices needed, and the subroutine PRINT is used to print the adjacency matrix with zeros suppressed. Output from the program includes titling information, number of variables, the adjacency matrix without zeros, and the strong and weak components. One member components are not printed.

The cards required to compute the components of the example problem of Chapter II using CLUSTR are

```
00007
TEST ADJACENCY MATRIX
(10X,7I2)
ADJ 1    0 1 0 0 0 0 0
ADJ 2    1 0 1 0 0 0 0
ADJ 3    0 0 0 0 0 0 0
ADJ 4    0 0 0 0 0 0 0
ADJ 5    0 1 1 0 0 1 0
ADJ 6    0 0 0 0 0 0 1
ADJ 7    0 0 0 0 1 0 0
```

Output for this example is found after the program in Appendix B.

CLUSTR uses 26,400 words of memory in the GE625 with 10,000 words in block common. A drum or disc is used for temporary storage during computation. Time required is less than .05 hours. The program accepts up to 100 variables.

CHAPTER IV: AN APPLICATION OF CLUSTR

In order to illustrate the use of CLUSTR in an actual study, it was applied to some data collected by the author in the spring of 1968 on the antecedents of temporary depression. The study, in the form of a questionnaire given to students in beginning psychology classes who were experiencing a temporary depression, was based on several unpublished studies by Shelly (1968) which indicated that people seek an optimal arousal level and are happiest with doing or experiencing things which contribute to reaching that optimal level. For instance, to a bored person arousing events such as listening to loud music, meeting new people, and engaging in a controversial discussion may be very pleasing to him; for someone who has had a particularly hectic day, listening to soft music in his favorite easy chair may be what pleases him most. It was conjectured that students in the midst of a temporary depression might have experienced events just before their depression which had caused their arousal levels to be less than ideal. Questions were asked about what sorts of things had happened in the twelve hours preceding the temporary depression and about what sorts of things they liked in general. The questionnaire may be found in Appendix C. The actual data is in Appendix D. There were 52 questions and 53 subjects.

Similarity matrices of two types were computed -- a regular Pearson r correlation matrix and one based on the omega-squared statistic[1]. Adjacency

1. Population index of the relative reduction in the variance of Y given the X value for an observation:

$$\omega^2 = \frac{\sigma_Y^2 - \sigma_{Y|X}^2}{\sigma_Y^2} \qquad \text{(Hays, 1963)}$$

matrices, corrected for directional relationships, were computed based on .3, .4, and .5 cutoff levels for the correlation matrix and on .09, .1, and .2 cutoff levels for the omega-square matrix and the variables were clustered using the program CLUSTR.

The clusters for the various levels of correlation are shown in Figure 4 and those for the levels of omega-square in Figure 5. Solid lines indicate the strong components and broken lines the weak components. The actual clusters may be found in Appendix B.

Among the patterns which can be seen from the various levels of clustering are two or three large clusters with several smaller clusters both within and without the larger clusters. One of the larger clusters, on the right in Figures 1 and 2 seems to include mainly general attitudes and preferences reflecting optimal arousal levels. Among the smaller clusters included is one concerned with the number and influence of close friends (40, 45) and another cluster concerned with liking dancing and controversial discussions (33,35). Another interesting cluster is the weak component found in the correlation clusters which links controversial discussions, change in opinion about Vietnam, and loudness of music (17,36,42).

A second large cluster (lower left) is concerned with events of the day preceding the temporary depression. Among the variables which seem to go together in this cluster are those of excitement, novel experiences, change in relationship with a friend, and number of pleasant things that happened (7, 13,5,15,24,26,4). A second group has to do with boredom, relaxation, amount of sleep, and number of unpleasant things (25,27,8,11,14). The first seems to deal with increasing arousal and the presence or absence of pleasant experiences. The second group deals more with decreasing arousal and presence or absence of unpleasant experiences. The relative independence of pleasant and unpleasant events is an observation also made by Shelly (1967).

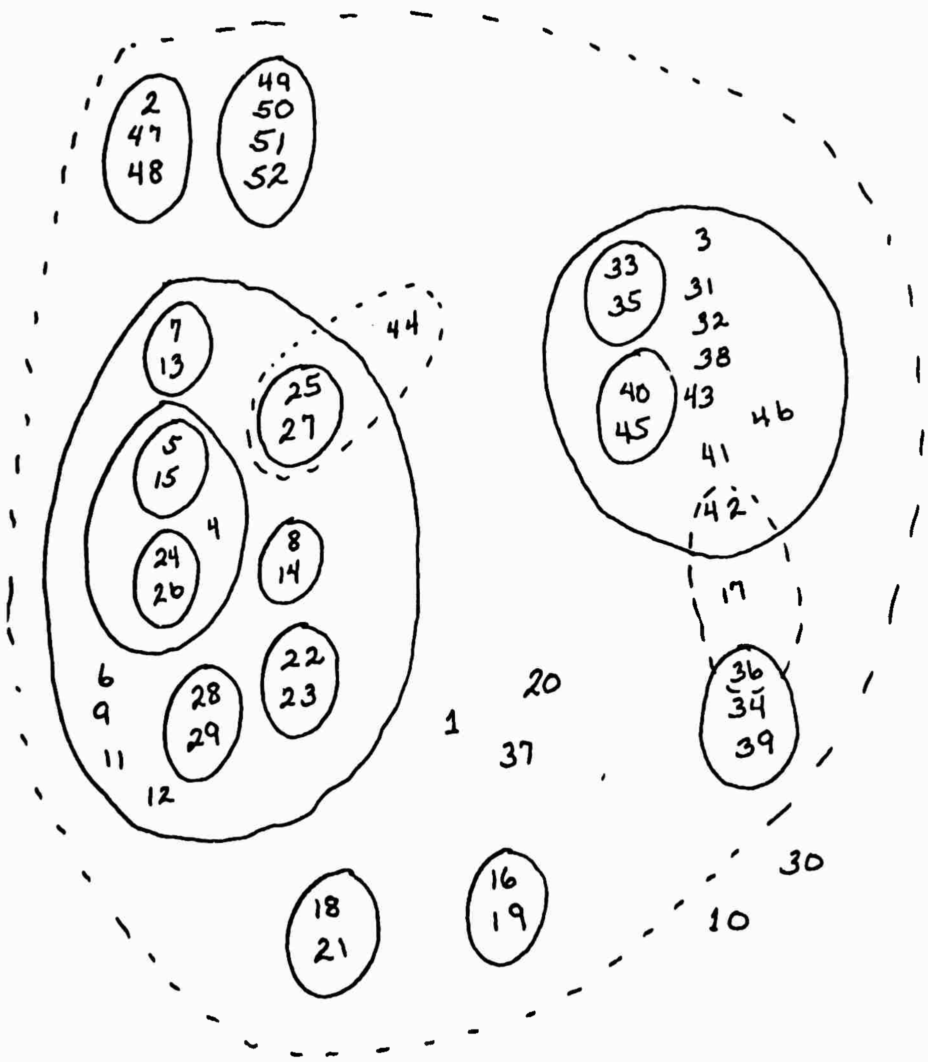


Figure 4. Clusters of the adjacency matrices based on cutoff levels of .3, .4, and .5 of the correlation matrix. Solid lines represent strong components, broken lines weak components. Innermost clusters represent the .5 cutoff level, outermost clusters represent the .3 cutoff level.

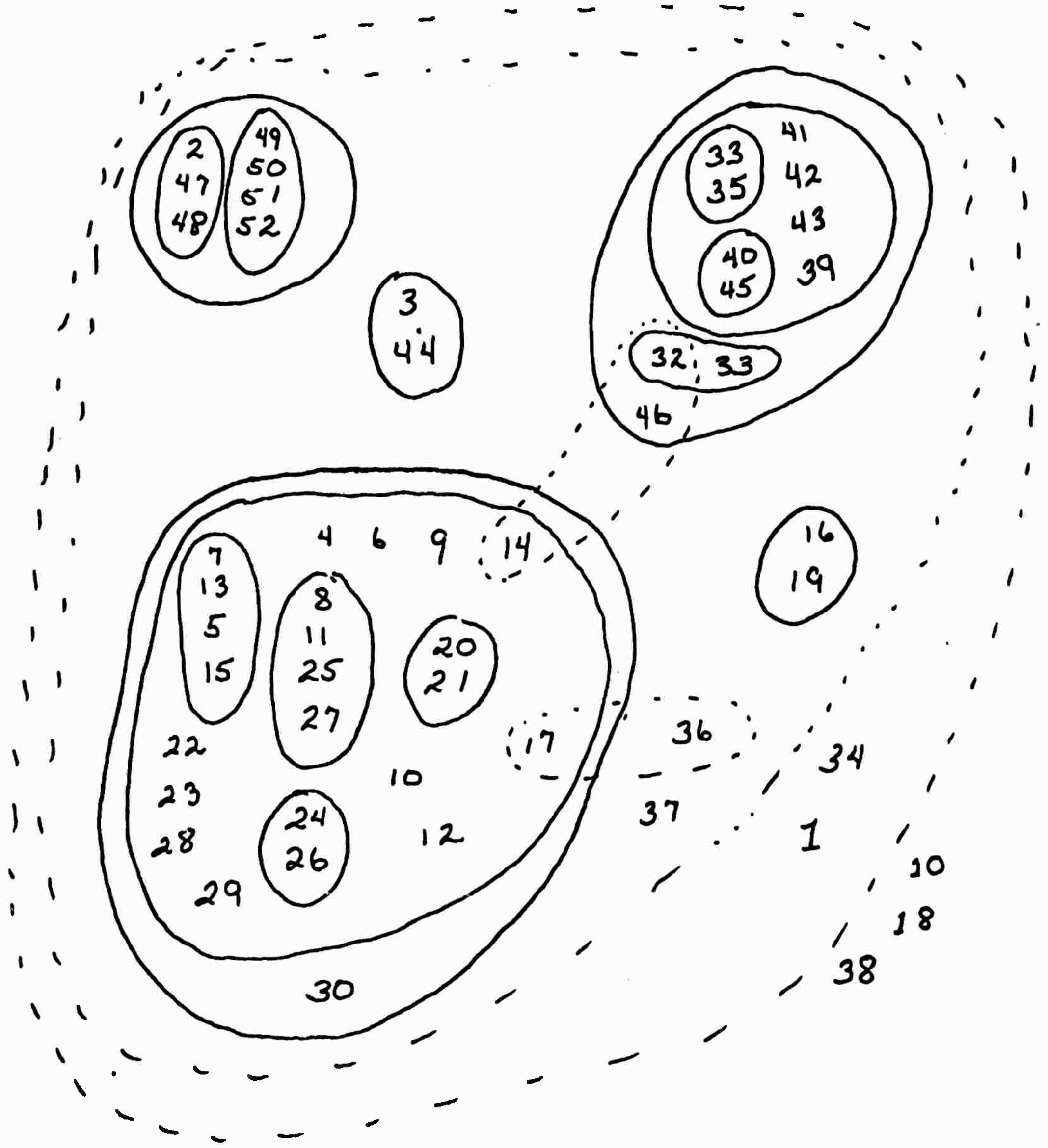


Figure 5. Clusters of the adjacency matrices based on cutoff levels of .09, .1, and .2 of the omega-squared matrix. Solid lines represent strong components, broken lines weak components. Innermost clusters are those based on the .2 cutoff level.

A third large cluster links severity, length, and course of the depression (2,47-52). The small cluster (16,19) which occurs in both groups of clusters seems to indicate that trouble with a friend is a separate reason for a temporary depression independent of other changes in arousal level.

A more thorough study of the clusters and the similarity matrices would yield information as well about the direction (in the sense of positive or negative correlations) and strength of relationships. It should be remembered that the difference between strong and weak components is not in the strength of relationships, but in the logical progression of the relationships. Relative strength of relationship can be discovered by comparing the clusters made at different cutoff levels, however. It can be seen that CLUSTR can provide quite useful information about the patterns of relationships which exist in the system as it is represented by the data.

BIBLIOGRAPHY

Ball, G. H. Data Analysis in the Social Sciences: What about the Details?, Proceedings Fall Joint Computer Conference, 1965, 533-59.

Ball, G. H. & Hall, D. J. ISODATA, a Novel Method of Data Analysis and Pattern Classification, Stanford Research Institute, Menlo Park, California, April 1965.

Bangert, C. J. Unpublished Computer Program: SFAIDB, University of Kansas Computer Center, 1968.

Cartwright, D. & Harary, F. Structural Balance: a Generalization of Heider's Theory, Psych. Rev., 1956, 63, 277-293.

Cattell, R. A Note on Correlation Clusters and Cluster Search Methods, Psychometrika, Vol 9, No. 3, September 1944.

Erdos, P. & Katona, G. Theory of Graphs, New York: Academic Press and Budapest: Publishing House of the 1st Hungarian Academy of Sciences.

Flament, C. Applications of Graph Theory to Group Structure, Englewood Cliffs, N. J.: Prentice-Hall, 1963.

Frank, Ronald E. & Green, Paul E. Numerical Taxonomy in Marketing Analysis, J. Marketing Research, V 1, February 1968, 83-94.

Harary, Frank & Norman, Robert Z. Graph Theory as a Mathematical Model in Social Science, Institute for Social Research, University of Michigan, Ann Arbor, Mich., 1953.

Harary, Frank; Norman, Robert Z. & Cartwright, Dorwin. Structural Models: an Introduction to the Theory of Directed Graphs, John Wiley & Sons, Inc., New York, 1965.

Hays, William L. Statistics for Psychologists, Holt, Rinehart, & Winston, New York, 1963.

Konig, D. Theorie der endlichen und unendlichen Graphen. Leipzig: Akademische Verlagsgesellshaft M. B. H., 1936.

Lindzey, G. & Gorgatta, E. F. Sociometric Measurement, in G. Lindzey (Ed.), Handbook of Social Psychology, Cambridge, Mass.: Addison - Wesley 1954, pp. 405-448.

Luce, R. D. & Raiffa, H. Games and Decisions, New York: Wiley, 1957.

Ore, O. Theory of Graphs. Providence: Amer. Math. Soc., Colloquims Publications, Vol 38, 1962.

Shelly, Maynard W. Unpublished studies, 1967.

Shelly, Maynard W. Unpublished manuscript, 1968.

Shelly, M. W.; McClelland, G. & Crow, B. A. DIGRAPH system, University of Kansas Computation Center, 1968.

Sokal, R. R. & Sneath, P. H. A. Prinicipl<u>es of Numerical Taxonomy</u>, W. H. Freeman & Co., San Francisco, 1963.

Tyron, R. C. & Bailey, D. E. The BCTRY Computer System of Cluster and Factor Analysis, Multivariate Behav. Res., 1, 1966, 95-111.

APPENDICES

APPENDIX A:

MATRIX AND BOOLEAN OPERATIONS

APPENDIX A

Matrix operations

Addition

$$C = A + B = (a_{ij} + b_{ij} = c_{ij})$$

Cross product (element-wise multiplication)

$$C = A \times B = (a_{ij}b_{ij} = c_{ij})$$

Multiplication - number of rows = number of columns = n

$$C = AB = (a_{i1}a_{1j} + a_{i2}b_{2j} + \dots) = (\sum_{k=1}^{n} a_{ik}b_{kj} = c_{ij})$$

Boolean arithmetic

$$\begin{array}{ll} 1 + 0 = 0 & 0 \,.\, 1 = 0 \\ 0 + 0 = 0 & 0 \,.\, 0 = 0 \\ 1 + 1 = 0 & 1 \,.\, 1 = 1 \end{array}$$

Exactly the same as regular arithmetic except 1 + 1 = 1.

Adjacency matrix

$$A = \begin{cases} a_{ij} = 1 \text{ if } v_i\, v_j \in \text{digraph} \\ a_{ij} = 0 \text{ if } v_i\, v_j \in \text{digraph} \end{cases}$$

Reachability matrix

$R = A + A^2 + A^3 + \dots$ with diagonal elements = 1.

Computation of reachability matrix from adjacency matrix.

Since there are seven points in the example problem, the greatest number of lines which are required to reach one point from another is six. Since the (n)th power of an adjacency matrix gives the paths of length n in the corresponding digraph, the greatest number of powers that will need to be taken in order to find all the paths in the digraph is six.

The entries of 1 in the original adjacency matrix A are the paths of length 1.

$$A = A^1 = \text{paths of length } 1 = \begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Multiplying A by itself we get A^2, the paths of length 2. The first two entries are

$$a_{11}^2 = a_{11}a_{11} + a_{12}a_{21} + a_{13}a_{31} + a_{14}a_{41} + a_{15}a_{51} + a_{16}a_{61} + a_{17}a_{71}$$

$$= 0 + 1 + 0 + 0 + 0 + 0 + 0 = 1$$

$$a_{12}^2 = a_{11}a_{12} + a_{12}a_{22} + a_{13}a_{32} + a_{14}a_{42} + a_{15}a_{52} + a_{16}a_{62} + a_{17}a_{72}$$

$$= 0 + 0 + 0 + 0 + 0 + 0 + 0 = 0$$

$$A^2 = \text{paths of length } 2 = \begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Continuing to take powers, we find

$$A^3 = \text{paths of length } 3 = \begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

$$A^4 = \text{paths of length } 4 = \begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

$$A^5 = \text{paths of length } 5 = \begin{array}{c|c} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Adding the powers using Boolean addition

$$A + A^2 + A^3 + A^4 = A + A^2 + A^3 + A^4 + A^5 = \begin{array}{c|c} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Since adding A^5 gives exactly the same matrix as before it was added, it is unnecessary to continue the process of taking powers and adding. When the diagonal elements are set to 1, the reachability matrix is complete.

$$R = \begin{array}{c|c} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

APPENDIX B:

THE PROGRAM CLUSTR AND SAMPLE OUTPUT

```
C       MAINLINE TO CALL CLUSTR
C
C     PARAMETER CARDS
C        1 COL 1-5 NVAR
C        2 TITLE
C        3 VARIABLE FORMAT
C
      DIMENSION IA(100,100), KTITLE(14),FMT(14)
      COMMON /Q/IA /R/NVAR, KTITLE
    1 FORMAT(I5)
    2 FORMAT(13A6,A2)
    8 READ 1,NVAR
      IF(NVAR.EQ.0) GO TO 9
      READ 2,KTITLE
      READ 2,FMT
      DO 3 I= 1,NVAR
    3 READ FMT, (IA(I,J),J = 1,NVAR)
      CALL CLUSTR
      GO TO 8
    9 STOP
      END
```

```
CCLUSTR     COMPUTATION OF STRONG AND WEAK COMPONENTS OF DIAGRAPH
C             WRITTEN BY BARBARA CROW      JUNE, 1968
C
C    THIS PROGRAM COMPUTES AND LISTS THE MEMBERS OF THE STRONG AND
C    WEAK COMPONENTS OF A DIGRAPH.
C    INPUT IS AN ADJACENCY MATRIX.
C
     SUBROUTINE CLUSTR
     DIMENSION IA(100,100),       WEAK(100,100),KTITLE(14),
    1FMT(3),FMS(3),ICOMP(100),NVEC(100),IR(100,100),STRONG(100,100)
     COMMON /Q/IA /R/NVAR,KTITLE
     INTEGER WEAK,STRONG,FMT,FMS,FMR
     EQUIVALENCE (WEAK,STRONG,IR)
     DATA(FMT(I), I =1,3)/6H(10X, ,6H000000,6HI4)    /
     DATA FMR/6HX,0000/
     DATA NW/07777/
     DATA (NVEC(I), I - 1,100)/100*0/
     DATA((IR(I,J), J = 1,100),I = 1,100)/10000*0/
C          PRINTING TITLE AND INPUT MATRIX
     CALL PRINT(IA,KTITLE,NVAR,0)
C                       COMPUTING STRONG COMPONENTS
C             COMPUTE REACHABILITY MATRIX OF ORIGINAL DIGRAPH
     CALL REPUN(NVAR,IR)
C         MULTIPLY ELEMENT WISE  IR AND ITS TRANSPOSE
     DO 50 I =1,NVAR
     DO 50 J =1,NVAR
     STRONG(I,J) = IR(I,J)*IR(J,I)
  50 CONTINUE
C        PRINT STRONG COMPONENTS
     NCP = 0
     WRITE(6,51)
  51 FORMAT(//20X,36HTHE STRONG COMPONENTS OF THE DIGRAPH//)
C       SET UP VECTOR OF ELEMENTS OF THE COMPONENT
     DO 55 I = 1,NVAR
     IF(NVEC(I).EQ.1) GO TO 55
C        INITIALIZE VECTOR ICOMP AND K
     DO 56 L = 1,NVAR
  56 ICOMP(L) = 0
     K = 0
C        FORM VECTORS OF ELEMENTS OF COMPONENTS
     DO 57  J = 1,NVAR
     IF(STRONG(I,J),EQ.0)  GO TO 57
     K  = K + 1
     ICOMP(K) = J
     NVEC(J) = 1
  57 CONTINUE
C         ELIMINATE ONE MEMBER COMPONENTS
     IF(K.EQ.1) GO TO 55
C         SET UP PART OF FORMAT
     NCP = NCP + 1
     FMS(1) = FMT(1)
     FMS(3) = FMT(3)
```

```
C        FINISH FORMAT
      IF(K.LT.10)  GO TO 58
C     TWO DIGIT FIELD PROBLEM
      NQ = NUMB(K,6)
      NG = AND(NW,NQ)
      FMS(2) = FMR = NG
      GO TO 59
   58 FMS(2) = FMR + K
C        PRINT COMPONENT
   59 WRITE(6,42) NCP
   42 FORMAT(1H0,9X,29HVARIABLES IN STRONG COMPONENT,I3)
C           SET UP LINE TO PRINT
      IS = MINO(K,30)
      DO 13 J1 - 1,K,IS
      J2 = MINO(J1+IS-1,K)
      WRITE(6,FMS) (ICOMP(J),J = J1,J2)
   13 CONTINUE
   55 CONTINUE
C              COMPUTING WEAK COMPONENTS
C        INITIALIZE NVEC,ISUM,WEAK
      DO 30 I = 1,NVAR
   30 WEAK(I,J) = 0
      FMS (2) = 0
C       ADD IA AND ITS TRANSPOSE TO FORM SYMMETRICIZED DIGRAPH
      DO 10 I = 1,NVAR
      DO 10 J = 1,NVAR
      IA(I,J)  = IA(I,J) + IA(J,I)
      IF(IA(I,J).GT.1) IA(I,J) = 1
   10 CONTINUE
C        COMPUTE REACHABILITY MATRIX OF SYMETRICIZED DIGRAPH
      CALL REFUN(NVAR,WEAK)
C        PRINT OUT WEAK COMPONENTS
      NCP = 0
      WRITE(6,16)
   16 FORMAT(//20X,34HTHE WEAK COMPONENTS OF THE DIGRAPH//)
C        SET UP VECTOR OF MEMBERS OF THE COMPONENT
C        ENTRIES OF 1 IN ROW GIVE MEMBERS OF THE SAME COMPONENT
      DO 22 I = 1,NVAR
      IF(NVEC(I).EQ.1) GO TO 22
C       INITIALIZE VECTOR ICOMP AND K
      DO 19 L = 1,NVAR
   19 ICOMP(L) = 0
      K = 0
C       FORM VECTORS OF MEMBERS OF EACH COMPONENT
      DO 20 J = 1,NVAR
      IF(WEAK(I,J).EQ.1) GO TO 20
C          K IS THE NUMBER OF ELEMENTS IN A COMPONENT
      K = K + 1
      ICOMP(K) = J
      NVEC(J) = 1
   20 CONTINUE
C        ELIMINATE ONE MEMBER COMPONENTS
      IF(K.EQ.1) GO TO 22
```

```
C        FINISH FORMAT
      NCP = NCP + 1
      IF(K.LT.10) GO TO 25
C        ELIMINATE PROBLEMS OF TWO DIGIT FIELD WIDTH
      NQ = NUMB(K,6)
      NG = AND(NW,NQ)
      FMS(2) = FMT(2) + NG
      GO TO 26
   25 FMS(2) = FMT(2) + K
C          PRINT COMPONENT
   26 WRITE(6,44) NCP
   44 FORMAT(1H0,9X,27HVARIABLES IN WEAK COMPONENT,I3)
C          SET UP LINE TO PRINT
      IS = MINO(K,30)
      DO 12 J1 = 1,K,IS
      J2 = MINO(J1+IS-1,K)
      WRITE(6,FMS) (ICOMP(J),J = J1,J2)
   12 CONTINUE
   22 CONTINUE
      RETURN
      END
```

```
CREPUN       CALCULATION OF REACHABILITY OR N-REACHABILITY MATRIX
C
C    THIS PROGRAM CALCULATES A REACHABILITY MATRIX FROM AN ADJACENCY
C    MATRIX.  THE VALUE OF N FOR AN N-REACHABILITY MATRIX MAY BE
C    SPECIFIED.  REACHABILITY MATRIX MAY BE PRINTED OR PUNCHED.
C
     SUBROUTINE REPUN(NIT,IJ)
     DIMENSION IA(100,100),IJ(100,100),IQ(100,100),KTITLE(14)
     COMMON /A/IA  /R/NVAR,KTITLE
     REWIND 1
     DO 3   I = 1,NVAR
     DO 3   J = 1,NVAR
C      ABSOLUTE VALUE OF IA
     IA(I,J) = IABS(IA(I,J))
C INITIALIZE IJ AND IQ AND DRUM
     IQ(I,J) = 0
  3  IJ(I,J) = IA(I,J)
     WRITE (1)  ((IA(I,J),J = 1,NVAR),I = 1,NVAR)
     REWIND 1
C      L IS THE NUMBER OF BOOLEAN POWERS TAKEN
     L = 0
     NT = NIT-1
     DO 9   K = 1,NT
     L = L + 1
C      COMPUTE BOOLEAN POWERS OF IA
     DO 5   I = 1,NVAR
     DO 5   J = 1,NVAR
     DO 4 M = 1,NVAR
  4  IQ(I,J) = IQ(I,J) + IA(I,M)*IJ(M,J)
  5  IF(IQ(I,J).GT.1)  IQ(I,J) = 1
C       READ IN REACHABILITY MATRIX FROM DRUM
     DO 1   I = 1,NVAR
     DO 1   J = 1,NVAR
  1  IJ(I,J) = 0
     READ (1) ((IJ(I,J),J = 1,NVAR), I = 1,NVAR)
     REWIND 1
C      CHECK FOR CONVERGENCE
     DO 6   I = 1,NVAR
     DO 6   J = 1,NVAR
     IF(IQ(I,J).EQ.1.AND.IJ(I,J).EQ.0)  GO TO 7
  6  CONTINUE
     GO TO 12
C      SUM POWERS OF IA
  7  DO 2   I = 1,NVAR
     DO 2   J = 1,NVAR
     IJ(I,J) = IJ(I,J) + IQ(I,J)
  2  IF(IJ(I,J).GT.1) IJ(I,J) = 1
C    CHECK FOR COMPLETED N-REACHABILITY MATRIX
     IF(K.EQ.NT) GO TO 12
C    STORE MATRIX ON DRUM
     WRITE (1) ((IJ(I,J),J = 1,NVAR), I = 1,NVAR)
     REWIND 1
```

```
C    REINITIALIZE IJ AND IQ
     DO 8   I = 1,NVAR
     DO 8   J = 1,NVAR
     IJ(I,J) = 0
     IJ(I,J) = IQ(I,J)
  8  IQ(I,J) = 0
  9  CONTINUE
C       SET DIAGONALS EQUAL TO 1
 12  DO 13  I = 1,NVAR
 13  IJ(I,I) = 1
     RETURN
     END
```

```
CPRINT
C
C     THIS SUBROUTINE SUPRESSES NUMBER LABELED NBLK AND PRINTS OUT
C     MATRIX WITH BLANKS IN THOSE SPACES.  I1 ONLY
      SUBROUTINE PRINT(IAJ,KTITLE,NVAR,NBLK)
      DIMENSION IAJ(100,1),KTITLE(14),FORM(105)
      INTEGER ALEFT,ARIGHT,ASPEC,BLANKS,X3SPEC,ASPECO,FORM
      DATA ALEFT,ARIGHT,ASPEC,ISPECM,BLANKS,X3SPEC,ISPEC,ISPECO,ASPECO/
     16H(     ,6H)      ,6HA1,    ,6HI1,    ,6H        ,6H3X,    ,6HI3,    ,
     26HI1    ,6HA1    /
C     SET UP OPENING AND CLOSING PARENTHESES FOR OUTPUT FORMAT
      FORM(1) = ALEFT
      FORM(NVAR + 5) = ARIGHT
          (2) = X3SPEC
      FORM(3) = ISPEC
      FORM(4) = X3SPEC
    7 FORMAT(1H1, 25X, 13A6, A2)
      WRITE(6,7) (KTITLE(I), I=1,14)
    8 FORMAT (1H0,25X,21HNUMBER OF VARIABLES = I3)
      WRITE(6,8)NVAR
    5 FORMAT(1H0,8X,1H1,8X,2H10,8X,2H20,8X,2H30,8X,2H40,8X,2H60,8X,
     12H60,8X,2H70,8X,2H80,8X,2H90,8X,3H100//)
      WRITE(6,5)
      DO 30  I = 1,NVAR
      MVAR = NVAR-1
      DO 31  J = 1,MVAR
      IF(IAJ(I,J).EQ.NBLK)  GO TO 32
      FORM(J+4) = ISPECM
      GO TO 31
   32 FORM(J+4) ASPEC
      IAJ(I,J) = BLANKS
   31 CONTINUE
      IF (IAJ(I,NVAR).EQ.NBLK)  GO TO 29
      FORM(NVAR+4) = ISPECO
      GO TO 28
   29 FORM(NVAR+4) = ASPECO
      IAJ(I,NVAR) = BLANKS
C     WRITE OUT ARRAY UNDER CONTROL OF COMPUTED FORMAT
   28 WRITE(6,FORM)  I, (IAJ(I,J),J = 1,NVAR)
      DO 11  J = 1,NVAR
   11 IF(IAJ(I,J).EQ.BLANKS)  IAJ(I,J) = 0
   30 CONTINUE
      RETURN
      END
```

TEST ADJACENCY MATRIX

NUMBER OF VARIABLES = 7

```
         10        20        30        40        50        60

1          1
2         1 1
3
4
5          11  1
6               1
7             1
```

THE STRONG COMPONENTS OF THE DIGRAPH

```
VARIABLES IN STRONG COMPONENT    1
   1   2

VARIABLES IN STRONG COMPONENT    2
   5   6   7
```

THE WEAK COMPONENTS OF THE DIGRAPH

```
VARIABLES IN WEAK COMPONENT    1
   1   2   3   5   6   7
```

APPENDIX C:

QUESTIONNAIRE

ANTECEDENTS OF TEMPORARY DEPRESSION

Instructions

Everyone occasionally goes through temporary depressions. This study is an investigation of some of the factors that may be associated with developing a temporary depression. Try to be as careful as possible in filling out the attached questionnaire.

You are to keep the attached questionnaire until, as happens to almost everyone, you go through a temporary depression. There is no need to return this questionnaire quickly; if you go through the rest of the semester without a temporary depression you will still receive one credit. If you do experience a temporary depression, as you are likely to, then fill out the attached questionnaire sometime during the depression.

After you have completed the questionnaire bring it to the Social Psychology Secretary in Room 607 Fraser. The secretary will give you one credit for your participation.

When the results have been analysed (late this summer or early fall) you may pick up a copy of the results in the Social Psychology Office.

Thank you very much for your cooperation.

ANTECEDENTS OF TEMPORARY DEPRESSION QUESTIONNAIRE

1. Sex

    1. male

    2. female

2. How long has it been since this temporary depression began? (Use the time you think it began, not when you first noticed it.)

    1. less than two hours

    2. two to four hours

    3. four to six hours

    4. one day

    5. two days

    6. three days

    7. four days to a week

    8. more than a week

3. What time of day did this temporary depression begin?

    1. morning

    2. afternoon

    3. evening

    4. late at night

Most of the following questions refer to the twelve waking hours preceding the temporary depression:

4. How happy were you in the twelve waking hours preceding this temporary depression?

1. not as happy as usual
2. about as happy as usual
3. a little happier than usual
4. much happier than usual

5. Did you do anything in the twelve waking hours preceding the depression that you usually don't do?

1. no, nothing unusual
2. something somewhat unusual
3. something quite unusual

6. Did you meet anyone you hadn't met before in these twelve preceding hours?

1. yes
2. no

7. How exciting were the twelve waking hours preceding the temporary depression?

1	2	3	4	5	6	7	8	9
extremely dull								extremely exciting

8. How relaxing were they?

1	2	3	4	5	6	7	8	9
extremely relaxing								extremely tense

9. How often did you engage in arguments (friendly or otherwise) in the

twelve waking hours preceding the depression?

1. not at all
2. once
3. twice
4. three times
5. more than three

10. How much did you have to do that day?

1. very little
2. not too much
3. quite a bit
4. a great deal to do

11. How much sleep did you get the night preceding the temporary depression?

1. much less than usual
2. a little less than usual
3. about the same as usual
4. a little more than usual
5. a lot more than usual

12. How many times did you feel really relaxed in the twelve waking hours preceding the depression?

1. not at all
2. once
3. twice
4. three times
5. more than three times

13. How many times did you feel very excited?

1. not at all
2. once
3. twice
4. three times
5. more than three times

14. How many times did you feel really bored?

1. not at all
2. once
3. twice
4. three times
5. more than three times

15. Did your relationship to one of your friends change for the better in the twelve waking hours preceding the temporary depression?

1. no
2. yes, somewhat
3. yes, our relationship changed a great deal for the better

16. Did your relationship with a friend change for the worse?

1. no
2. yes, somewhat
3. yes, a great deal for the worse

17. How many times did you engage in controversial discussions in the twelve waking hours preceding the depression?

1. not at all

2. once

3. twice

4. three times

5. more than three times

18. How many different places did you go in the twelve hours preceding the depression?

1. one to two

2. three to four

3. five to ten

4. ten to twenty

5. more than twenty

19. During the preceding twelve hours did a person of the opposite sex cause you much trouble?

1. none

2. a little

3. quite a bit

4. a great deal

20. Did you do better than expected on something you did in the twelve hours preceding the depression (or receive the results of something you had done earlier)?

0. doesn't apply

1. no, no better than expected

2. somewhat better than expected

3. much better than expected

21. Did you do worse than expected on something?

  0. doesn't apply
  1. no, no worse than expected
  2. somewhat worse than expected
  3. much worse than expected

22. How many people were you with during most of the twelve waking hours preceding the temporary depression?

  1. none, I was alone most of the time
  2. one or two other people
  3. three to five other people
  4. five to ten
  5. ten to twenty
  6. more than twenty

23. The number of people you were with was

  1. less than usual
  2. about the same as usual
  3. more than usual

24. How many pleasant things happened to you in the twelve waking hours preceding the depression?

  1. none
  2. a few
  3. some
  4. quite a few
  5. many

25. How many unpleasant things happened?

1. none

2. few

3. some

4. quite a few

5. many

26. Was the number of pleasant things greater or smaller than usual?

1. less than usual

2. about same as usual

3. more than usual

27. Was the number of unpleasant things greater or smaller than usual?

1. less than usual

2. about same as usual

3. more than usual

28. When the temporary depression began to come on were you

1. in a very familiar place

2. in a somewhat familiar place

3. in a place you had never been before

29. How many associations does this place arouse for you?

1	2	3	4	5	6	7	8	9
very few								a great many associations

30. Were there any specific events which you feel precipitated the temporary depression?

1. don't know
2. nothing in particular
3. a series of events
4. a single event

The following are very general questions about life of a college student at KU.

31. To what extent are you "going with" someone of the opposite sex?
    1. not at all
    2. dating one person more than any other
    3. dating only one person
    4. engaged
    5. married

32. How often do you go places to meet new people?
    1. several times a day
    2. several times a week
    3. once a week
    4. once every two weeks
    5. less than every two weeks

33. To what extent do you enjoy discussing controversial subjects?
    1. extremely well
    2. quite well
    3. all right once in a while
    4. not very well
    5. not at all

34. How happy is your life at KU in general?

1	2	3	4	5	6	7	8	9
miserable								extremely happy

35. Do you enjoy dancing?

1. yes
2. sometimes
3. no

36. Do you prefer loud or soft music?

1. loud
2. in between
3. soft

37. How much do unfriendly arguments upset you?

1. not at all
2. very little
3. not too much
4. quite a bit
5. a great deal

38. To what extent do you enjoy working on problems you are unlikely to be able to solve?

1. a great deal
2. once in a while, or until I decide it can't be solved easily
3. I'd rather stick to problems I can solve

39. Do you dislike movies with tragic endings?

1. yes

2. no

40. How many close friends do you have?

1. none

2. one

3. two

4. three to five

5. five to ten

6. more than ten

41. Has your opinion about the draft changed in the last few months?

1. yes

2. no

42. About the Vietnam war?

1. yes

2. no

43. Do you prefer parties where there are

1. just two people

2. a few people

3. 10 to 20 people

4. more than 30 people

44. How far ahead do you like to make dates?

1. on the spur of the moment

2. one day

3. two days
4. a week
5. two weeks
6. more than two weeks

45. When you make decisions do you think about what close friends might do?

1. yes
2. no

46. How much do you enjoy a good meal?

1. a great deal, one of my favorite things to do
2. quite a bit
3. okay
4. just something that has to be done before going on to other things

To be answered as soon as possible after the temporary depression is gone.

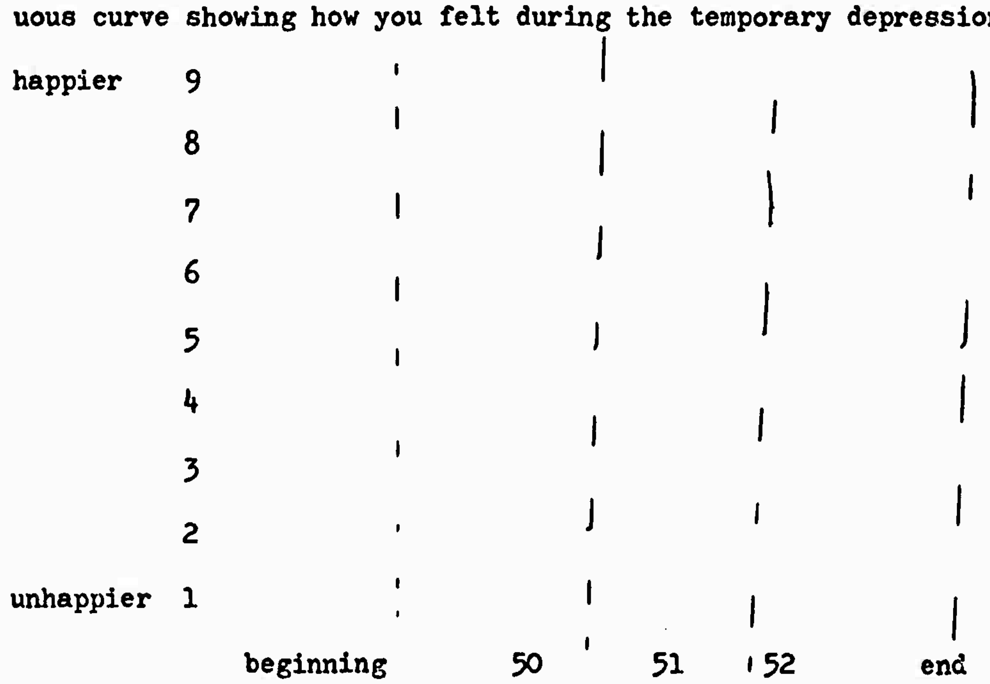
47. How long did the temporary depression last all together?

1. less than two hours
2. two - four hours
3. four - six hours
4. a day
5. two days
6. three days
7. four days - week
8. two weeks
9. more than two weeks

48. Compared with other temporary depressions you have experienced, how severe was this one?

1. less severe than most
2. about same
3. more severe than most

49. - 52. To help describe the course of the temporary depression, please use the chart below. Start at the left side of the chart and draw a continuous curve showing how you felt during the temporary depression.



APPENDIX D:

DATA, ADJACENCY MATRICES AND CLUSTERS

ANTE TEMP DEPR 01	2433329714325532211112223121833534224223223312427653
ANTE TEMP DEPR 02	1444123222252112113112222233133527214323221412416535
ANTE TEMP DEPR 03	2634319314155131131115351311743218135316223112725885
ANTE TEMP DEPR 04	2632124613333411111322233232245535325312222311722113
ANTE TEMP DEPR 05	2322125543223311521233212121742128123324113412313335
ANTE TEMP DEPR 06	2233217713112111111113221321741218124225223311314578
ANTE TEMP DEPR 07	222212452234231153124211212152122111532421341 2336
ANTE TEMP DEPR 08	14121257333112112111 22223221524556 3 4122221 23527321
ANTE TEMP DEPR 09	214212561333121151113621223194353412521 52 3312111115
ANTE TEMP DEPR 10	2112213412422211111236341311343437124313224222323236
ANTE TEMP DEPR 11	3432125724211211142134222131513237133325113412611237
ANTE TEMP DEPR 12	1133216511235112232232232221743517233225112111122113
ANTE TEMP DEPR 13	24143296223252313321152423222223235125225214312532215
ANTE TEMP DEPR 14	12331258122232212422222323327335533242224222323318557
ANTE TEMP DEPR 15	12311235343314111222342231317 2317123225222421215345
ANTE TEMP DEPR 16	2632126623224112231236322131432238113226214411736545
ANTE TEMP DEPR 17	2831217512125113133115324333931324124223113312831216
ANTE TEMP DEPR 18	1332225611511511212232223221515414123225122312425323
ANTE TEMP DEPR 19	1422211413312522121243233232432517234125222413418657
ANTE TEMP DEPR 20	2432124614211211121316221221711144334214222422624346
ANTE TEMP DEPR 21	2534127523225111213112244332642217135325212413532427
ANTE TEMP DEPR 22	1731126824311511331325232221933337224225214312725744
ANTE TEMP DEPR 23	1132113531323411133242222311931336324315123111323247
ANTE TEMP DEPR 24	1533214223311211132236232321723316231125212112632324
ANTE TEMP DEPR 25	1324125612223411132322222222333327224225213412614346
ANTE TEMP DEPR 26	2522114412311213121112212121933327234222212322737534
ANTE TEMP DEPR 27	2232226623324312313114222231532225114224214412423774
ANTE TEMP DEPR 28	1412125623224111232313233221533227223326223111422235
ANTE TEMP DEPR 29	1842125824112111323113212121614417124326112411515334
ANTE TEMP DEPR 30	2511321914121521113112224131343424115223223212422565
ANTE TEMP DEPR 31	2423127532254221222733612222432326114225113411214567
ANTE TEMP DEPR 32	2823134613332111132146341222141545333321222121924457
ANTE TEMP DEPR 33	18421176233323132131122332319322131 2124113422834326
ANTE TEMP DEPR 34	24211254 22321111211332231315333325324224221622325412
ANTE TEMP DEPR 35	2113127514113212143136234131912528124325113312311126
ANTE TEMP DEPR 36	15432264133551122323222432221333371243251234 1434312
ANTE TEMP DEPR 37	2212123511132311321325222221911136114326223411214567
ANTE TEMP DEPR 38	1622127412332111131241122222443335234215222422732467
ANTE TEMP DEPR 39	1841125953111511131242214131932513134314222412833322
ANTE TEMP DEPR 40	1431125622222212322225223211643234113315114412414122
ANTE TEMP DEPR 41	2223225513351111133244222221541537222223224422115336
ANTE TEMP DEPR 42	1232124613414211122223222221634527133225223312336433
ANTE TEMP DEPR 43	2122225424412121231322232222541328132215222311126337
ANTE TEMP DEPR 44	1512124513332413314333222221331223111326221121428621
ANTE TEMP DEPR 45	1224127213453112123344242312541224333113222312631123
ANTE TEMP DEPR 46	2423225622421111211315221221115546223114222412631237
ANTE TEMP DEPR 47	1142116432332211331225232222441226113121114422225434
ANTE TEMP DEPR 48	253212644443 22212411 222211922218122224123111525357
ANTE TEMP DEPR 49	28221268341535221312452422216245281342 4222323837546
ANTE TEMP DEPR 50	1112127654225111231115322223211417324224214421414789
ANTE TEMP DEPR 51	182312554231123121242224332193432713122211432173
ANTE TEMP DEPR 52	2323111933222511321231224231631514214221211422434226
ANTE TEMP DEPR 53	113321681311541212211332222113222712512422342111

ADJACENCY MATRICES

ANTECEDENTS OF TEMPORARY DEPRESSION   CORR.   .3

NUMBER OF VARIABLES = 52

```
      1        10        20        30        40        50        60

 1                                                       1
 2                                                  11
 3                                            1
 4           1     1         1 1
 5             1     1
 6                          1  1
 7        1        111
 8               11 1           11
 9        1
10
11            1
12            1
13       11  1              11  1
14          1
15         1                 1
16                      1
17
18                         1
19                     1
20
21                     1
22                           1  1
23         1      1
24       1           1         1
25                   1      1
26       1 1                1  11
27           1                11
28                  1          1  1
29                                1
30
31                                    1                   1
32                                   1   1    1  1
33         1                           1
34                                        1
35                                    11      11 1 1
36                      11    1              1
37                1 1
38                                                 11
39                               1        1
40                                    1  1         1
41                        1                     1       1
42                      1                      1 1     1
43            1             1         1  1      1
44                            1 11
45                                       1  1 1
46
47                                                    1
48                                                   1
49                                                      11
50                                                     1 1
51                                                     11 1
52                                                       1
```

ANTECEDENTS OF TEMPORARY DEPRESSION   CORR.   .4

NUMBER OF VARIABLES = 52

```
      1        10        20        30        40        50        60
  1
  2                                                 11
  3
  4                         1 1
  5                 1
  6
  7              1
  8               1
  9
 10
 11
 12
 13         1
 14          1
 15       1                 1
 16                     1
 17
 18
 19                    1
 20
 21
 22                        1
 23                       1
 24      1            1      1
 25                            1
 26      1                 1
 27                         1
 28                             1
 29                            1
 30
 31
 32
 33                                  1
 34
 35                                1
 36                   1
 37
 38
 39
 40                                            1
 41
 42                   1
 43
 44                         1
 45                                        1
 46
 47   1                                          1
 48   1                                         1
 49                                                1
 50                                               1 1
 51                                                1 1
 52                                                 1
```

ANTECENDENTS OF TEMPORARY DEPRESSION   CORR.   .5

NUMBER OF VARIABLES = 52

```
   1        10        20        30        40        50        60

 1
 2                                                1
 3
 4
 5                1
 6
 7              1
 8
 9
10
11
12
13        1
14
15      1
16
17
18
19
20
21
22
23
24                           1
25                            1
26                         1
27                          1
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47  1                                              1
48                                                1
49                                                   1
50                                                  1 1
51                                                   1 1
52                                                    1
```

ANTECEDENTS OF TEMPORARY DEPRESSION OMEGA. .09

NUMBER OF VARIABLES = 52

	1									10										20										30										40										50										60
1																																																				1								
2																																														1	1												
3																																											1																
4								1					1											1	1	1		1																																
5						1			1						1																																													
6					1																					1																																		
7													1	1	1																																													
8				1							1	1		1										1	1																																			
9					1																																													1										
10																																																												
11								1							1								1		1					1																		1												
12								1																																																				
13				1			1														1		1					1																																
14							1	1																	1			1	1																															
15					1		1				1													1																									1											
16																			1																																									
17																								1																																				
18																																																												
19																1																																												
20																					1																																							
21													1							1																																								
22																							1																																					
23											1		1									1																																						
24				1				1							1		1									1																						1												
25				1				1			1			1													1																								1	1								
26				1	1																			1			1																																	
27																									1	1																																		
28				1									1	1															1																															
29														1														1																																
30											1																																																	
31																																1																												
32														1																	1												1																	
33						1																													1				1							1														
34																							1																																					
35																																	1								1		1																	
36												1					1							1																							1	1												
37									1										1	1																													1											
38																																																												
39																											1			1			1									1																		
40																																													1															
41									1										1																1							1								1										
42									1								1								1														1		1		1																	
43																						1										1			1							1			1															
44			1																1					1		1	1																																	
45																																								1			1						1											
46														1																			1																1											
47		1																																														1												
48		1																																													1					1								
49																																																		1										
50																																																	1		1									
51																																																		1		1								
52																																																1			1									

ANTECEDENTS OF TEMPORARY DEPRESSION   OMEGA.   .1

NUMBER OF VARIABLES = 52

```
   1        10        20        30        40        50        60

 1
 2                                               11
 3                                            1
 4       1               111
 5     1  1     1
 6    1                    1
 7             111
 8   1      11 1         11
 9    1
10               1
11       1                  1
12       1    1
13      1    1        1 1    1
14      11
15    1 1  1             1
16                   1
17                       1
18
19                1
20                     1
21             1      1
22                       1
23             1        1
24   1   1      1 1        1                     1
25   1   1                  1                       11
26   11                   1
27           1             1
28             1               1
29                           1
30
31                               1
32              1                1
33                                  1   1
34
35           1                     1         1
36            1    1      1                      1
37        1 1       11
38
39                          1  1  1        1
40                                            1
41        1         1                      1       1
42        1 1     1       1             1 1 1
43                      1            1      1  1
44  1                1      11
45                                       1  1     1
46              1                                 1
47 1                                             1
48 1                                            1
49                                                 1
50                                                1 1
51                                                 1 1
52                                                  1
```

ANTECEDENTS OF TEMPORARY DEPRESSION OMEGA. .2

NUMBER OF VARIABLES = 52

```
    1        10        20        30        40        50        60

 1  
 2                                               1
 3  
 4  
 5               1
 6  
 7             1 1
 8           1             1
 9  
10  
11        1
12  
13       1
14  
15     1 1
16  
17  
18  
19  
20                     1
21                    1
22  
23  
24                          1
25        1                  1
26                        1
27  
28  
29  
30  
31  
32              1
33                                   1
34  
35                                 1
36                 1
37  
38  
39                           1
40                                             1
41  
42  
43  
44  
45                                        1
46              1
47  1                                              1
48                                                1
49                                                   1
50                                                  1 1
51                                                   1 1
52                                                    1
```

CLUSTERS

COMPONENTS BASED ON .3 CUTOFF LEVEL OF CORRELATION MATRIX

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48

VARIABLES IN STRONG COMPONENT 2

3 31 32 33 35 38 40 41 42 43 45 46

VARIABLES IN STRONG COMPONENT 3

4 5 6 7 8 9 11 12 13 14 15 22 23 24 25 26 27 28 29

VARIABLES IN STRONG COMPONENT 4

16 19

VARIABLES IN STRONG COMPONENT 5

18 21

VARIABLES IN STRONG COMPONENT 6

34 36 39

VARIABLES IN STRONG COMPONENT 7

49 50 51 52

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21

22 23 34 25 26 27 28 29 31 32 33 34 35 36 37 38 39 40

41 42 43 44 45 46 47 48 49 50 51 52

COMPONENTS BASED ON .4 CUTOFF LEVEL OF CORRELATION MATRIX

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48

VARIABLES IN STRONG COMPONENT 2

4 5 15 24 26

VARIABLES IN STRONG COMPONENT 3

7 13

VARIABLES IN STRONG COMPONENT 4

8 14

VARIABLES IN STRONG COMPONENT 5

16 19

VARIABLES IN STRONG COMPONENT 6

22 23

VARIABLES IN STRONG COMPONENT 7

25 27

VARIABLES IN STRONG COMPONENT 8

28 29

VARIABLES IN STRONG COMPONENT 9

33 35

VARIABLES IN STRONG COMPONENT 10

40 45

VARIABLES IN STRONG COMPONENT 11

49 50 51 52

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

2 47 48

VARIABLES IN WEAK COMPONENT 2

4 5 15 24 26

VARIABLES IN WEAK COMPONENT 3

7 13

VARIABLES IN WEAK COMPONENT 4

8 14

VARIABLES IN WEAK COMPONENT 5

16 19

VARIABLES IN WEAK COMPONENT 6

17 36 42

VARIABLES IN WEAK COMPONENT 7

22 23

VARIABLES IN WEAK COMPONENT 8

25 27 44

VARIABLES IN WEAK COMPONENT 9

28 29

VARIABLES IN WEAK COMPONENT 10

33 35

VARIABLES IN WEAK COMPONENT 11

40 45

VARIABLES IN WEAK COMPONENT 12

49 50 51 52

COMPONENTS BASED ON .5 CUTOFF LEVEL FOR CORRELATIONS

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48

VARIABLES IN STRONG COMPONENT 2

5 15

VARIABLES IN STRONG COMPONENT 3

7 13

VARIABLES IN STRONG COMPONENT 4

24 26

VARIABLES IN STRONG COMPONENT 5

25 27

VARIABLES IN STRONG COMPONENT 6

49 50 51 52

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

2 47 48

VARIABLES IN WEAK COMPONENT 2

5 15

VARIABLES IN WEAK COMPONENT 3

7 13

VARIABLES IN WEAK COMPONENT 4

24 26

VARIABLES IN WEAK COMPONENT 5

25 27

VARIABLES IN WEAK COMPONENT 6

49 50 51 52

COMPONENTS BASED ON .09 CUTOFF LEVEL FOR OMEGA-SQUARED

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48 49 50 51 52

VARIABLES IN STRONG COMPONENT 2

3 44

VARIABLES IN STRONG COMPONENT 3

4 5 6 7 8 9 11 12 13 14 15 17 20 21 22 23 24 25 26
27 28 29 30

VARIABLES IN STRONG COMPONENT 4

16 19

VARIABLES IN STRONG COMPONENT 5

31 32 33 35 39 40 41 42 43 45 46

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 19 20 21 22
23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 39 40 41
42 43 44 45 46 47 48 49 50 51 52

COMPONENTS BASED ON .1 CUTOFF LEVEL FOR OMEGA-SQUARES

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48

VARIABLES IN STRONG COMPONENT 2

3 44

VARIABLES IN STRONG COMPONENT 3

4 5 6 7 8 9 10 11 12 13 14 15 17 20 21 22 23 24 25
26 27 28 29

VARIABLES IN STRONG COMPONENT 4

16 19

VARIABLES IN STRONG COMPONENT 5

31 32

VARIABLES IN STRONG COMPONENT 6

33 35 39 40 41 42 43 45

VARIABLES IN STRONG COMPONENT 7

49 50 51 52

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 19 20 21 22
23 24 25 26 27 28 29 30 31 32 33 35 36 37 39 40 41 42
43 44 45 46 47 48 49 50 51 52

COMPONENTS BASED ON .2 CUTOFF LEVEL FOR OMEGA-SQUARES

THE STRONG COMPONENTS OF THE DIGRAPH

VARIABLES IN STRONG COMPONENT 1

2 47 48

VARIABLES IN STRONG COMPONENT 2

5 7 13 15

VARIABLES IN STRONG COMPONENT 3

8 11 25 27

VARIABLES IN STRONG COMPONENT 4

20 21

VARIABLES IN STRONG COMPONENT 5

24 26

VARIABLES IN STRONG COMPONENT 6

33 35

VARIABLES IN STRONG COMPONENT 7

40 45

VARIABLES IN STRONG COMPONENT 8

49 50 51 52

THE WEAK COMPONENTS OF THE DIGRAPH

VARIABLES IN WEAK COMPONENT 1

2 47 48

VARIABLES IN WEAK COMPONENT 2

5 7 13 15

VARIABLES IN WEAK COMPONENT 3

8 11 25 27 39

VARIABLES IN WEAK COMPONENT 4

14 32 46

VARIABLES IN WEAK COMPONENT 5

17 36

VARIABLES IN WEAK COMPONENT 6

20 21

VARIABLES IN WEAK COMPONENT 7

24 26

VARIABLES IN WEAK COMPONENT 8

33 35

VARIABLES IN WEAK COMPONENT 9

40 45

VARIABLES IN WEAK COMPONENT 10

49 50 51 52

REPORT #1
Digraph
Cluster
Strong Component
Weak Component

UNCLASSIFIED
Security Classification

AD

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Kansas Department of Anthropology Lawrence, Kansas 66044	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

A GRAPH THEORETICAL APPROACH TO CLUSTERING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

Barbara Ann Crow

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
Oct. 1968	70 pages	21

8a. CONTRACT OR GRANT NO. F44620-68-C-0081 (ARPA) b. PROJECT NO. 9543 c. 6154501D d. 681313	9a. ORIGINATOR'S REPORT NUMBER(S) #1 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR 69-0912TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research 1400 Wilson Boulevard Arlington, Virginia 22209 (SRLB)

13. ABSTRACT

This report describes a method of clustering variables based upon graph theory. The clusters computed are the strong and weak components of a digraph. This program was developed because of the absense of suitable programs for clustering variables when all of the relationships between the variables were not symmetric. The major part of the report describes the program, the computer system of which it is a part, and gives an example as to how it works together with sample output from the program.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification